

《心理学报》审稿意见与作者回应

题目：运动特征对抽象动词具身表征的影响：来自 fMRI 及 EMG 的证据

作者：李想，贾丽娜，魏士琳，陈俊涛，夏尧远，王芹，金花

第一轮

审稿人 1 意见：

本研究探讨了词语运动特征与抽象动词具身表征之间的因果联系，实验设计科学严谨，论证逻辑清楚，对于因以往研究结果分歧引发的对抽象概念表征具身认知观的质疑具有重要的意义。问题与建议如下：

意见 1：核磁测试后的行为实验的目的是什么？与核磁实验是否有关联？建议写清楚。

回应：非常感谢审稿专家严谨而中肯的建议！核磁测试后的行为实验的目的是为了同时对检验词语具身性常用的行为学方法（即 ACE）进行验证。行为测试本身与核磁实验没有关联，为了减轻被试的负担，两者均使用了真假词判断任务。

根据审稿专家的建议，我们在修改稿中第 2 部分的第 1 自然段补充了行为测试的目的等内容。

意见 2：行为测试没有发现显著的 ACE 结果，是否可能被试并未记牢新造词的意思，也就无法感知新造词的运动特征？

回应：非常感谢审稿专家严谨而中肯的意见和建议！对于行为测试没有获得显著 ACE 结果的原因，专家的推测“可能被试并未记牢新造词的意思，也就无法感知新造词的运动特征”是一种可能。但在本研究中，我们认为这种可能性不大。一方面，本研究实验第三天（第二次学习后）和第四天的“新造词-释义连线”纸笔测验正确率分别为 75.8% 及 74.3% 的结果，及核磁实验的结果即“两侧楔前叶在新造抽象动词的后测加工中存在的较强激活”均提示本研究中被试对新造词意义进行了成功习得及提取，支持了学习过程的有效性。另一方面，具身领域内已有的行为研究在具体和抽象语言加工中的 ACE 并未获得一致结果（如在具体和抽象语言中均观察到 ACE: Glenberg & Kaschak, 2002; Glenberg et al., 2008; 在具体语言中观察到 ACE: van Dam & Desai, 2017; Schaller et al., 2017; 王斌等, 2019; 在抽象语言中观察到 ACE: Guan et al., 2013; 在具体和抽象语言中未重复出 ACE 的: Morey et al., 2021; 在抽象语言中未得到 ACE 的: van Dam & Desai, 2017 等），这可能是由于 ACE 本身的不稳定性造成的。包括 ACE 在内的系列行为具身效应均采用了以语言-动作意义相符时的语义判断反应时显著低于不相符条件这一相符促进现象来推断运动系统在语言加工过程中的参与，但是如修改稿中引言第 3 自然段中提到的那样，有研究者也提出包括 ACE 在内的这种行为相符效应也可能出现在概念系统之外的反应阶段（Kiefer & Harpaintner, 2020），这可能在一定程度上解释了 ACE 等行为具身效应不稳定的原因。再者，虽然 ACE 在早期一直被认为是支持具身

语言理解观点的最有代表性的行为证据之一,但近年来的一些行为研究未能重复得到该效应并对其提出了质疑(如 Morey et al., 2021),对此,Zwaan(2021)在综述中也提到,那些未能重复得到如 ACE 等具身效应的研究,并不能作为反对运动系统在语言理解中参与的证据,而只能作为对这种效应及范式的一种质疑。因此有可能并不是所有的具身效应都能在行为层面观察到 ACE,如本研究的行为测试结果。相比而言,磁共振成像的方法可能更为敏感和客观。因此,我们推测“行为测试没有发现显著的 ACE 结果”可能与 ACE 自身结果的不稳定性有关。

根据审稿专家的建议,我们将纸笔测验的正确率结果补充在了修改稿中 2.3.1 部分的第 2 自然段;并且在讨论 4.1 部分的第 4 自然段增加了对这方面的说明。同时我们在修改稿中 2.4 讨论部分的第 3 自然段对行为测试未能发现显著的 ACE 结果的原因补充了详细解释。

参考文献:

王斌,李智睿,伍丽梅,张积家.(2019).具身模拟在汉语肢体动词理解中的作用.《心理学报》,51(12),1291-1305.

Glenberg, A. M., Sato, M., & Cattaneo, L. (2008). Use-induced motor plasticity affects the processing of abstract and concrete language. *Current Biology*, 18(7), 290-291.

Guan, C. Q., Meng, W., Yao, R., & Glenberg, A. M. (2013). The motor system contributes to comprehension of abstract language. *PLoS ONE*, 8(9), e75183.

Schaller, F., Weiss, S., & Muller, H. M. (2017). "Pushing the button while pushing the argument": motor priming of abstract action language. *Cognitive Science*, 41(5), 1328-1349.

van Dam, W. O., & Desai, R. H. (2017). Embodied simulations are modulated by sentential perspective. *Cognitive Science*, 41(6), 1613-1628.

Zwaan R. A. (2021). Two Challenges to "Embodied cognition" research and how to overcome them. *Journal of Cognition*, 4(1), 14.

意见 3: 第二次学习后被试纸笔测验的准确率多少? 建议写清楚,学习准确率对后续的实验会有很大影响。

回应: 非常感谢审稿专家严谨而中肯的意见和建议! 本研究实验第二天(第一次学习后)、实验第三天(第二次学习后)和第四天的“新造词-释义连线”测试正确率分别为 44.2%、75.8% 及 74.3%,实验第二、三天学习后的“圈出假词”测试的平均正确率分别为 89.2%和 95.8%。已将此部分的结果补充在了修改稿中的 2.3.1 部分的第 2 自然段。

意见 4:全文建议精简语言,突出重点。

回应: 非常感谢审稿专家的意见和建议! 我们对文章做了全面、仔细的修改,更正了文中的错别字、笔误等问题,对整个行文作了进一步的精简和凝练。

意见 5:讨论 4.3 标题建议用更明确、有内容的标题代替。

回应: 非常感谢审稿专家严谨而中肯的意见和建议! 根据专家的建议,基于具身认知理论的内容及本研究的实验结果,我们对这一部分的标题进行了完善和明确,已修改为“抽象语言理解的具身模拟程度依赖于词语运动特征的高低”。

.....

审稿人 2 意见:

《运动特征对抽象动词具身表征的影响: 来自 fMRI 及 EMG 的证据》一文通过两个实验, 探讨了运动特征信息影响抽象词汇加工的神经机制。结果发现了, 初级运动皮层以及高级中枢脑区在学习新造词后出现神经可塑性变化, 肯定了感知觉皮层在抽象概念加工中的重要作用, 支持了具身认知的语言理解观点。作者关注抽象动词的神经表征加工过程, 从新词学习视角考察抽象概念加工, 具有一定创新性和独特性。在方法上结合了学习-测试范式以及磁共振技术的高空间分辨率优点, 还利用了肌电图技术, 使该研究具有很强的研究必要性。但是文章在写作过程中还存在一些问题, 因此建议大修后再审。

主要有以下问题:

意见 1: 文章引言在内容组织上松散且宽泛、局部段落到整体内容的逻辑表述不清楚、缺乏实验间逻辑联系等问题。

回应: 非常感谢审稿专家严谨而中肯的意见和建议! 根据专家的意见和建议, 我们对问题提出部分的理论论述以及写作逻辑等方面做了全面的修改和梳理。对审稿专家提出的意见和建议进行了整理并逐一回答和说明如下:

1) 引言首段开篇详细介绍了双重编码理论以及传统认知观点主张, 但通读全文, 这些理论观点并非作者所表达支持的观点内容。建议作者认真思考后对内容进行重新整理, 突出介绍本研究重点关注的理论主张。

回应: 非常感谢审稿专家的指导和建议。根据审稿专家的建议, 基于本研究的关注重点, 我们对引言首段中的双重编码理论及传统认知观点等内容进行了删除, 并将原稿引言的前两个自然段内容合为一段。详见修改稿中引言第 1 自然段。

2) 研究背景的介绍比较单薄, 没有体现出从新词习得视角出发探究抽象词汇表征加工的重要性; 而且作者对于抽象具身表征结果分歧性的剖析尽管较为充分, 但在后面也并未提及控制抽象概念语义复杂性这一维度, 比如控制词汇视觉属性、情绪内省维度等。

回应: 非常感谢审稿专家的批评和指导! 根据审稿专家的指导意见和建议, 我们在修改稿中补充了相关文献和介绍, 同时对整个引言部分的行文逻辑进行了仔细的修改和梳理。

对于“从新词习得视角出发探讨抽象词汇表征加工的重要性”, 我们在修改稿中引言的第 3 自然段增加了对与新词习得有关的观点和研究的介绍与分析, 详见修改稿的引言部分第 3 自然段第 7 行“其中, Li 等(2022)的研究采用学习-测试范式...”至 19 行的内容。

对于“控制抽象概念语义复杂性”这一问题, 我们重新阅读了部分文献, 对这部分内容的逻辑及语言表达进行了仔细修改和梳理, 详见修改稿中引言的第 2 自然段。对于“词汇视觉属性、情绪内省维度”的控制等, 由于本研究实验均采用了重复测量的设计, 所以在实验中并未控制; 针对情绪维度可能产生的影响, 我们在原稿中讨论部分的不足与展望中已进行了说明。

3) 研究重点关注的“动作句子相符效应(ACE)”并未在引言部分作出具体解释说明,而是在讨论“4.3 理论探索”的第二段作出了具体解释。建议作者认真思考后,重新组织前言内容。

回应: 非常感谢审稿专家的批评和指导!我们在修改稿的引言部分第 3 自然段中补充了对“ACE”的具体解释,详见该段的第 3、4 行。同时结合引言部分存在的其他问题,我们对引言各部分内容的写作逻辑、文字表达等进行了全面的修改和仔细的梳理。

4) 引言部分第五段强调说明了运动特征对抽象动词在运动皮层内表征的影响有必要深入探究,但第六段又突然转入到了“对于抽象词语加工过程是否也存在外周肌肉运动系统的共振活动”。两段之间的逻辑关系作者并未详细展开说明。建议作者认真思考两个实验间的逻辑关系以及结合整体研究问题,对该部分进行详细说明。

回应: 非常感谢审稿专家的建议和意见!根据审稿专家的建议,我们在修改稿的引言部分第 4 自然段中补充说明了两个实验间的逻辑关系等内容,即该段前 4 行及后 3 行。

5) 作者基于金等人(2022)以及 Li 等人(2022)的两篇行为研究,对于 fMRI 实验一神经层面结果进行了预期,支持证据不够充分,建议补充相关文献依据。

回应: 非常感谢审稿专家的批评和指导!感觉审稿专家的建议,我们在修改稿中引言的第 5 自然段中补充了相关文献依据。同时结合引言部分存在的其他问题,我们对引言各部分内容的写作逻辑、文字表达等进行了全面的修改和仔细的梳理,并在修改稿中引言的第 2 和 3 自然段中增加了对相关文献的详细介绍和分析。

意见 2: 方法结果部分中,存在表述冗余、不准确等问题。

回应: 非常感谢审稿专家严谨而仔细的意见和建议!笔者对审稿专家提到的意见和建议进行了整理并逐一回答和说明如下:

1) 被试性别不平衡(男生比例仅占总人数的 10%左右),这可能对结果产生一定影响。

回应: 谢谢审稿专家的意见和建议!本研究为了避免被试的个体差异等因素可能对实验结果产生的影响,两个实验均采用了被试内实验设计。

但是,更严格地说,我们可能还是无法排除当前性别比率对结果外推的影响。Glenberg 等(2009)在对以往文献进行综述后提出,在具身语言理解的研究中,在情绪和权力维度上可能存在性别差异,如由于女性更容易对悲伤事件反应、男性更容易对愤怒事件反应,所以女性理解描述悲伤事件的句子时快于男性,而男性理解愤怒事件的句子时比女性更快。但至今未见有研究探讨语言理解过程中运动模拟的性别差异。

另一方面,关于性别对词汇语义加工的影响,研究结果上尚未达成一致。Xu 等(2020)使用语义判断任务(要求被试判断视觉呈现的 2 个词语是否为近义词)考察语言理解过程中的性别差异,结果,基于大脑激活水平和左右半球偏侧化的分析均没有发现性别差异;但基于功能连接的分析发现了大脑功能连接模式存在显著的性别差异,表现在语言关键脑区间的动态连接模式,以及男性表现出更强的局部信息分离,而女性表现出更强的全局信息整合。黎樱等(2011)在分析了相关文献后提出,语义加工性别差异的研究分歧可能源自不同研究采用了不同类型的语言任务或范式,有的实验设计可能并不只是反映了语言加工过程。本研究

的实验任务均为真假词判断，而 Hill 等(2006)研究中的真假词判断任务和 Ihnen 等(2009)的词语和非词阅读任务都未得出性别差异的结论。最近一些关注词汇语义加工的神经成像研究在采用了重复测量实验设计后也并未特意平衡被试的性别，如 Gao 等(2019)的 fMRI 研究的被试为 18 名被试（男性 6 名），王汉林等(2020)的两个 ERP 实验被试分别为 26（男生 7 人）和 27 人（男生 9 人），Kiefer 等(2022)的 ERP 研究的被试为 44 名（男性 8 名）等。同这些研究一样，考虑到性别因素及效应并非本研究重点考察和关注的内容，因此本研究在招募被试时并未对被试的性别等进行平衡。

参考文献：

- 黎樱, 杨东, 张庆林. (2011). 语言加工性别差异的神经机制. *心理科学进展*, 19(11), 1625-1634.
- 王汉林, 蒋泽亮, 冯晓慧, 鲁忠义. (2020). 道德概念的空间形象性:语言因素和具身因素的共同作用. *心理学报*, 52(2), 128-138.
- Gao, C., Baucom, L. B., Kim, J., Wang, J., Wedell, D. H., & Shinkareva, S. V. (2019). Distinguishing abstract from concrete concepts in supramodal brain regions. *Neuropsychologia*, 102-110.
- Glenberg, A. M., Webster, B. J., Mouilso, E., Havas, D., & Lindeman, L. M. (2009). Gender, emotion, and the embodiment of language comprehension. *Emotion Review*, 1(2), 151-161.
- Hill, H., Ott, F., Herbert, C., & Weisbrod, M. (2006). Response execution in lexical decision tasks obscures sex-specific lateralization effects in language processing: Evidence from event-related potential measures during word reading. *Cerebral Cortex*, 16, 978-989.
- Ihnen, S. K. Z., Church, J. A., Petersen, S. E., & Schlaggar, B. L. (2009). Lack of generalizability of sex differences in the fMRI BOLD activity associated with language processing in adults. *Neuroimage*, 45, 1020-1032.
- Kiefer, M., Pielke, L., & Trumpp, N. M. (2022). Differential temporo-spatial pattern of electrical brain activity during the processing of abstract concepts related to mental states and verbal associations. *NeuroImage*, 252, 119036.
- Xu, M., Liang, X., Ou, J., Li, H., Luo, Y. J., & Tan, L. H. (2020). Sex differences in functional brain networks for language. *Cerebral Cortex*, 30(3), 1528-1537.

2) 实验材料选定部分描述过于冗长繁琐。例如，正文第 5 页第二段，“同 Li 等人(2022) 研究中的实验 2 一样...”，后面又赘述了相应的描述内容。再者，下一段内容中，作者将研究目的与材料筛选杂糅在一起，会让读者产生一定混乱。建议精简整理材料筛选过程。

回应：非常感谢审稿专家的批评和指导！根据审稿专家的指导意见和建议，我们对 2.1.2 实验材料的选定与描述进行了全面修改和精简。

对于原稿中 2.1.2 实验材料最后一个自然段，为了避免带给读者混乱，我们将这部分关于实验材料中的抽象动词运动特征评分结果及评定目的等内容分别放在了修改稿中第 2 部分 2.1.2 实验材料部分的第 3 自然段最后 3 行以及 2.2.2 统计分析部分的第 2 自然段中的 7-9 行。

3) 作者未对笔画数、部件数、词频等心理语言学变量进行匹配，请对材料进行统计，如果没有匹配，需要阐述这些因素对结果的解释是否存在潜在的风险。

回应：非常感谢审稿专家的批评和指导！

根据审稿专家的建议，我们将实验 1 中新造词的选取方式及笔画数等内容补充在了 2.1.2 实验材料的介绍中，详见修改稿中该部分第二自然段中的 1-4 行。

其中，学习新造词（120 个汉字）和非学习新造词（120 个汉字）的平均笔画数分别为 7.69 和 7.32($t=1.159, p=0.248$)、平均字频的对数值分别为 3.85 和 3.73($t=1.318, p=0.189$)。所有新造词均由 2 个汉字随机组成，均通过了无意义评定，可以认为这些新造词的词频为 0。参照以往新词有关的研究（如白学军等，2019），我们单独对新造词和非新造词的所有汉字字频和笔画数进行了统计，结果显示两个条件中的汉字字频和笔画数差异均不显著。

对于心理语言学变量可能产生的影响，本研究实验 1 采用的是 2×2 两因素重复测量设计，主要关注的是对于运动特征增加后的学习新造词被试在后测加工和前测加工中观察到的神经激活差异，因此只对新词所表示释义的运动特征及方向性进行了评定和匹配；而且如前所述，学习和非学习新造词之间的笔画数和字频不存在显著差异，所以实验 1 的结果不会受到字笔画数、字频等的影响。

实验 2 采用单因素重复测量设计，词语类型包含了学习新造词、具体动词和非学习新造词，这三种类型词语的平均笔画数分别为 7.77±2.47、9.12±2.88 和 7.72±2.89，三者的字频对数平均值分别为 4.00±0.84、3.59±0.80 和 4.02±0.68。实验 2 主要关注的是学习新造词和作为基线的非学习新造词之间的外周肌肉活动差异，这两种类型词语的平均笔画数和字频对数平均值均不存在显著差异 ($ps>0.05$)，因此并不对结果构成影响。而具体动词在上述两个变量的平均值均显著高于另外两种类型 ($ps<0.05$)，但考虑到具体动词加工的具身效应已在神经成像研究中基本达成一致（本文引言第 1 自然段也有提及），且具体动词加工能够引发外周肌肉运动系统的共振激活这一结果也已在多数研究中得到过证实（如 Fino et al., 2016; Gálvez-García et al., 2020），因此推测上述变量也不会对实验 2 结果产生影响。

对于实验 2 未匹配实验材料的笔画数等变量这一局限，我们在不足与展望部分补充了相关内容，并会在后续研究中使用匹配更加严格的实验材料对本研究结果作进一步验证和探究。补充内容详见修改稿的讨论部分的不足与展望。

参考文献：

白学军, 马杰, 李馨, 连坤予, 谭珂, 杨宇等. (2019). 发展性阅读障碍儿童的新词习得及其改善. *心理学报*, 51(4), 471-483.

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PloS one*, 5(6), e10729.

4) 实验一行为数据结果不完整，建议补充。例如，作者仅提到了“被试均认真完成了新造词学习和记忆过程”，未说明正确率统计结果。

回应：谢谢审稿专家的意见和建议！本研究实验第二天（第一次学习后）、实验第三天（第二次学习后）和第四天的“新造词-释义连线”测试正确率分别为 44.2%、75.8% 及 74.3%，实验第二、三天学习后的“圈出假词”测试的平均正确率分别为 89.2% 和 95.8%，这些都表明了被试对新造词的意义进行了有效学习。已将此部分的结果补充在了修改稿中的 2.3.1 部分的第 2 自然段。

5)结果图需要进一步规范和美观。比如，建议调整图 2、3 阈值，以显示不同脑区激活强弱变化；建议增加柱状图中的误差线。

回应：谢谢审稿专家的意见和建议！根据审稿专家的指导意见和建议，在图 3 中各条件在左侧中央前回的 beta 值柱状图已添加了平均数 ± 1 个标准误的误差线，并对图 2 的脑激活图片进行了修改。

6) 建议补充实验 2 中的 t 检验对应的效果量。

回应：谢谢审稿专家的意见和建议！根据审稿专家的指导意见和建议，已在修改稿中实验 2 的 3.3 结果部分的第一自然段补充了 t 检验对应的效果量等信息。

意见 3：文章讨论部分不够深入，对于结果解释不够全面

回应：非常感谢审稿专家严谨而仔细的意见和建议！笔者对审稿专家在该部分提到的意见和建议进行了整理并逐一回答和说明如下：

1) 在实验 1、2 的分讨论中，作者仅详细叙述了实验重要结果，并未结合相应的文献依据，对结果做出充分讨论。但作者详细说明了在实验 1 中行为后测任务中未能得到学习新造词的显著 ACE 结果的原因以及实验 2 中指屈肌的肌电活动未出现预期结果的原因。这很大程度上忽略了讨论的重点内容。建议作者对主要结果进行详细解释，突出本文的亮点内容，而不是重点解释与之前研究结果存在出入。

回应：谢谢审稿专家的指导和帮助！根据审稿专家的建议，我们对实验 1 和实验 2 各自分讨论部分的内容进行了全面的修改和仔细的梳理，详见修改稿的 2.4 讨论和 3.4 讨论部分。

2) 作者所列举的文献证据对本研究结果支持性不足。本文利用汉语双字词为学习材料，侧重探讨抽象运动属性概念语义神经表征过程，但作者在解释学习新造词和非学习新造词的前、后测对比结果时，引用了来自拼音文字的研究证据，提出了被试在前测加工汉语新造词更依赖于语音加工。但与拼音文字相比，汉语作为表意文字，字形到语音的对应关系较弱，这些支持证据似乎并不能很好解释本研究所发现的颞中回、颞上回以及额上回等脑区的激活，建议作者认真思考该部分结果，补充相关汉语加工文献，对此作出合理解释。

回应：非常感谢审稿专家的指导和帮助！对于该问题，我们查阅文献发现，关于汉语加工的神经基础，目前的研究结果尚未达成一致。一些学者认为尽管字母文字和表意文字的书写有很大差异，但是其神经机制是一致的，即跨语言阅读的普遍性；另一种观点认为，由于文字间的差异，文字系统的性质影响了不同脑区的加工（汪洁等 2021）。两种观点各得到了不同的研究支持。本研究使用的是双字假词，汉语双字词通常作为一个整体被加工（隋雪，张晓利，2012），Tan 和 Perfetti 等（1999）的脑成像研究使用意义判断任务发现汉语双字词的识别中语音激活早于语义激活。本研究的结果一定程度上支持了语音激活在汉语双字词识别和加工中的作用。

根据审稿专家的指导意见和建议，我们在修改稿中 2.4 讨论部分的第 2 自然段最后 5 行，针对解释学习新造词和非学习新造词的前、后测对比结果，补充了相关文献及说明。

参考文献：

隋雪, 张晓利. (2012). 汉语双字词识别研究进展. *辽宁师范大学学报: 社会科学版*, 35(6), 768-771.

汪洁, 宋为群, 吴东宇, 袁英. (2021). 汉字阅读中语音加工的神经机制. *中国康复医学杂志*, 36(10), 1315-1321.

Tan, L. H. , & Perfetti, C. A. . (1999). Phonological activation in visual identification of Chinese two-character words. *Journal of Experimental Psychology Learning Memory & Cognition*, 25(2), 382-393.

3) 在总讨论的部分中存在论据罗列, 内容逻辑缺乏等问题。例如, “4.1 词语运动特征影响抽象动词的具身表征”的第二段以及最后一段文字内容, 都在罗列前人文献中发现了双侧/左侧中央前、后回或者角回的发现, 并未细致展开对于本文研究结果的讨论。此外, 作者所推论出的一些总结性语句也未结合具体结果来进行说明。例如, “实验 1 中左侧角回的较强激活, 既表明了新造词抽象意义的成功习得, 也可能反映了抽象概念基于特定模态、异极模态等脑区的多种模态表征方式”, 建议作者在结合相关文献证据基础上, 对本研究所发现的具体结果表现模式作出较为充分讨论, 进而得出对应的小结。

回应: 非常感谢审稿专家的指导和建议! 根据审稿专家的指导意见和建议, 我们对 4.1 讨论部分的行文逻辑和语言表述等进行了全面的修改和仔细的梳理。

意见 4: 对于文章结论第一条存在疑问。本研究结果可以得出感知觉运动皮层参与表征抽象概念特征信息, 但似乎无法从神经层面结果中得出运动特征在抽象动词的具身表征中发挥因果作用。建议作者后续结合 TMS、tDCS 等神经调控技术来抑制运动皮层, 进一步考察被试学习新造抽象动词后所诱发的运动脑区激活变化, 进而验证运动特征在抽象动词的具身表征中发挥因果作用。

回应: 非常感谢审稿专家严谨而仔细的建议和指导! 这为我们今后的研究提供了很好的方向, 已有研究者也曾提出 TMS 是探究因果联系的重要神经科学技术(Ostarek & Huettig, 2019), 在后续研究中我们会结合使用 TMS、tDCS 等对因果关系验证更具说服力的技术来进一步探究词语运动特征在新造抽象动词具身表征中的因果作用。我们将这部分内容补充在了修改稿中的讨论部分。同时对结论部分的第一条表述作出了修改。

意见 5: 一些细节问题:

回应: 非常感谢审稿专家严谨而仔细的意见和建议! 笔者对审稿专家指出的细节问题进行了逐一修改, 同时对全文进行了反复检查和校对。具体的修改说明如下:

1) 图 1 标题错误, “实验 5”应改为“实验 1”

回应: 非常感谢审稿专家的意见和建议! 已对此处错误进行修改, 并且对全文进行了反复校对和修改。

2) 文中表格以图片呈现, 且表 1、2 中因线条颜色粗细不一致, 引起表格数据不清晰, 建议以规范的三线表格式呈现在正文中。

回应: 非常感谢审稿专家的意见和建议! 根据审稿专家的建议, 我们已将文中所有表格按照更规范的格式在正文中进行了重新插入和编辑, 详见修改稿中 2.3.2 部分的所有表格。

3) 对于功能性磁共振名称应统一，文中核磁、fMRI、磁共振成像同时出现，建议统一术语。
回应：非常感谢审稿专家的指导和建议！根据审稿专家的建议，已对文章做了全面、仔细的修改和检查，对与功能磁共振名称有关的表述进行了统一。

4) 文中有出现漏字现象（例如，“2.1.5 神经成像数据的获取”中所写的“单次激发梯度回波（EPI）平面成技术”）
回应：非常感谢审稿专家的意见和建议！根据专家的建议和意见，已对文章做了全面、仔细的修改，更正了文中的错别字、笔误、标点符号和逻辑连词等诸如此类的问题。

5) 建议去除像“这样”、“呢”等相对口语化词汇，让文中语句更加严谨以及精炼。
回应：非常感谢审稿专家的意见和建议！根据专家的建议和意见，已对文章的语言表达做了全面、仔细的修改，删除和修改了文中的口语化词汇和不规范表达。

6) 存在语句不通顺问题，例如，“这可能表明了位于临近前运动皮层的额上回...功能”，建议通读全文，修改类似问题。
回应：非常感谢审稿专家的意见和建议！根据专家的建议和意见，已对此处内容进行了修改。同时对文章中存在的其他语法问题进行了全面检查和仔细修改，对整个行文也作了进一步的精简和凝练。

.....

审稿人 3 意见:

研究《运动特征对抽象动词具身表征的影响：来自 fMRI 及 EMG 的证据》采用功能磁共振成像技术（fMRI）以及表面肌电图技术考察抽象词汇的具身表征神经基础，对于推动具身认知的研究具有重要价值。然而，研究的实验设计中存在有待澄清的问题，引言逻辑也有待完善。具体如下：

意见 1：引言逻辑。

回应：非常感谢审稿专家严谨而仔细的意见和建议！笔者对引言部分的逻辑进行了重新梳理，对审稿专家提到的意见和建议进行整理并逐一回答和说明如下：

1) 引言第三段分析“有关抽象概念具身表征结果的分歧性可能和抽象概念的语义复杂性有关。”然而，整段的证据都是抽象和具体概念中感觉运动经验参与程度不同，和“抽象概念的语义复杂性”这一讨论主题偏离。

回应：非常感谢审稿专家的意见和建议！根据审稿专家的指导意见和建议，我们结合引言部分存在的其他问题，对包括本段在内的引言各部分内容的写作逻辑和文字表达等进行了全面的修改和仔细的梳理，详细修改内容见修改稿中引言部分的第 2 自然段。

2) 引言第四段, “那些未能在抽象词语加工中观察到感觉运动脑区激活的研究(如 Willems et al., 2011; Repetto et al., 2013; Innocenti et al., 2014), 可能使用了感觉运动经验权重较低(即各种感觉特征或运动特征较低)的词语材料”, 请列举这些文献中使用的词汇, 更直接地说明“可能使用了感觉运动经验权重较低”这一点。此外, 为何此处引用的研究和前文中列举的文献“但确实有些研究并没有发现感觉运动有关的脑区在抽象概念加工中的激活(Grossman et al., 2002; Dalla Volta et al., 2014; Innocenti et al., 2004; Tomasino et al., 2018)”不一致?

根据这一逻辑, 那么在抽象词语加工中观察到感觉运动脑区激活的研究(如 Mosley et al., 2012; Muraki et al., 2020; Harpaintner et al., 2020; Dreyer & Pulvermüller, 2018; Kiefer et al., 2022)使用的词汇是“感觉运动经验权重较高”的吗? 请补充例子。

回应: 非常感谢审稿专家的意见和建议! 下面对该问题中包含的三个小问题进行逐一回答和说明。

首先, 根据审稿专家的建议, 已在文中引言部分补充了那些未能在抽象词语加工中观察到感觉运动脑区激活的研究使用的抽象词语示例, 以“更直接地说明“可能使用了感觉运动经验权重较低”这一点”。详见修改稿中引言部分第2自然段。

其次, 对于文中列举的抽象概念加工引发的感觉运动激活阴性结果文献在前后文中的不一致问题, 原稿引言第一段中引用的文献(如 Grossman et al., 2002; Dalla Volta et al., 2014; Innocenti et al., 2014; Tomasino et al., 2018)均为未能发现感觉运动脑区在抽象概念加工中激活的支持证据, 经过文献阅读与分析, 发现其中的部分文献结果可能是由于使用了感觉运动经验权重较低的抽象词语, 因此仅列举了这几篇文献, 但忽略了文献引用前后一致性的问题。感谢审稿专家的意见, 已进行了相应修改和补充, 同时检查了全文, 保证前后文献引用的一致。此外, 文中引用一处文献年份出现笔误, 已将(Innocenti et al., 2004)更正为(Innocenti et al., 2014); 同时仔细检查了全文, 避免同类错误的出现。

最后, 对被认为使用了感觉运动经验权重较高词语的研究(如 Mosley et al., 2012; Muraki et al., 2020; Harpaintner et al., 2020; Dreyer & Pulvermüller, 2018; Kiefer et al., 2022), 列举其中的词语示例如下。如 Dreyer 和 Pulvermüller (2018) 研究中的实验材料为情绪含义抽象名词(如“fear”)和心理含义抽象名词(如“logic”)。Moseley 等(2012)则根据词语的具体性评分、可表象性评分以及动作关联性的评分, 得到了与感觉运动联系强弱不同的感觉运动抽象情绪动词(如“wail”)和纯抽象情绪动词(如“dread”)。Harpaintner 等(2020)在研究中基于前人的评定研究使用了高运动特征抽象名词“activity”和高视觉特征抽象名词“beauty”。Kiefer 等(2022)的 ERP 研究中使用的实验材料为心理特征抽象词如“surprise”和言语特征抽象词如“fate”。上述文献及示例已在修改稿中的引言和讨论的相应部分进行了补充。

意见 2: 实验设计逻辑。研究通过将人造词赋予意义来实现“运动特征增加”, 从而考察运动特征的增加影响抽象动词具身表征的神经基础。根据目前的实验设计, 前测中这些学习词汇对被试而言是“无意义的假词”, 学习后成为了“有意义的真词”。因此, 前后测差异可能只是由意义的增加引发的, 换言之, 因变量的变化可能是由“假词”到“真词”的改变引起的, 并非由“动词意义的增加”引起的。同样, 学习生造词和非学习生造词的差异也是由“真词”和“假词”的差异引起。如果要检验“运动特征的增加对抽象动词具身表征的影响”这一问题, 至少应该增加一个控制条件(例如赋予某些假词非动词意义), 确保因变量的变化确实是由“运动特

征的增加”这一操纵引起。如果实验设计中的这一问题确实存在，那么本研究得到的结果则无法解决“运动特征的增加影响抽象动词具身表征的神经基础”这一根本问题。请作者解释。

回应：非常感谢审稿专家严谨而仔细的意见和建议！

审稿专家提出的“如果要检验运动特征的增加对抽象动词具身表征的影响这一问题，至少应该增加一个控制条件（例如赋予某些假词非动词意义），确保因变量的变化确实是由运动特征的增加这一操纵引起。”这一指导建议给了我们一个很好的启发，增加一个赋予某些假词非动词意义的控制条件确实会使结果更为纯粹和表现出更绝对的说服力，我们将在后续研究中采用相应的实验范式和实验任务进一步探讨该问题。

对于专家提出的“因变量的变化可能是由假词到真词的改变引起的，并非由动词意义的增加引起的”这一问题，我们进行了认真思考，认为“因变量的变化可能是由动词意义的增加引起的”这样的解释还是成立的，主要解释如下：

如审稿人所言，前测中的“无意义的假词”在学习后成为了“有意义的真词”，这一过程确实属于广义上的语义增加过程而非独立的动词意义增加过程。但从具体的实验操作上考虑，可以认为本研究前测中的“无意义的假词”在学习后成为了“有意义的真抽象动词”，前后测间存在语义上的增加是“抽象动词意义增加”这一变化。在实验操作中，我们将表示抽象动作含义的动词短语描述作为学习新造词的释义，如“本绪：把责任推出去”。假词到真词的变化实际就是从无意义的新造词到有意义的抽象动词的变化；而且测试阶段中旨在检验被试的词语意义学习效果的“新造词-释义连线”纸笔测试结果（实验第二天即第一次学习后、实验第三天即第二次学习后和第四天即后测中被试在该项测试上的平均正确率分别为 44.2%、75.8%及 74.3%；这一结果已补充在修改稿中 2.3.1 部分的第 2 自然段），也从另一角度证明了被试在抽象动词意义上的学习效果，表明被试确实在学习新造词的学习过程中习得了抽象动词的词语意义并且表征了动作经验即“增加了词语运动特征”。

此外，实验 1 的参数调节分析得到的“词语运动特征值显著预测了后测中学习新造词在包括右侧中央前后回、左侧中央前回在内的两个较大团块的神经激活强度”这一结果，也验证了本研究实验操作的有效性，表明因变量的变化不只是假词到真词的变化引起的，而确实是由抽象动词的运动特征增加导致的。在已有的非动词含义即抽象名词有关的研究中，高运动特征抽象名词（如“activity”）(Harpaintner et al., 2020) 或与面部动作联系紧密的情绪抽象名词（如“fear”）(Dreyer 和 Pulvermüller, 2018) 等的加工均引发了运动脑区的参与，而那些未考虑运动特征差异或未评定运动特征的抽象名词加工研究多数并没有观察到运动脑区激活(如 Noppeney & Price, 2004; Sabsevitz et al., 2005)，基于这些研究也可推测，名词性或动词性的意义本身可能并非影响词语具身表征的关键，主要是包含运动特征的意义增加提高了运动系统在抽象词语加工中的参与度。缺乏赋予部分假词非动词意义的控制条件，使本研究无法从神经层面对比运动特征增加对抽象名词和抽象动词具身表征的影响的差异，但是我们认为采用目前的实验设计还是能够为考察“运动特征的增加影响抽象动词具身表征的神经基础”这一问题提供实验支持。

根据审稿专家的指导建议，我们在修改稿中的 4.1 讨论部分最后 1 个自然段中补充了相关内容。

参考文献：

Noppeney, U., & Price, C. J. (2004). Retrieval of abstract semantics. *NeuroImage*, 22(1), 164-170.

第二轮

审稿人 2 意见: 我的审稿意见得到了作者非常细致地修改和回复。没有更多意见或建议。建议接收发表。

审稿人 3 意见:

作者认真回答了审稿人的问题，并逐一进行修改和完善，回答及修改都较为合理，文章质量得到了较大幅度的提升。仍存在的问题是：

意见 1: 关于本人提出的因变量的变化可能是由“假词”到“真词”的改变引起的，并非由“动词意义的增加”引起的问题，虽然作者认为“从具体的实验操作上考虑，可以认为本研究前测中的“无意义的假词”在学习后成为了“有意义的真抽象动词”，前后测间存在语义上的增加是“抽象动词意义增加”这一变化”的解释有一定合理性，且“实验 1 的参数调节分析得到的“词语运动特征值显著预测了后测中学习新造词在包括右侧中央前后回、左侧中央前回在内的两个较大团块的神经激活强度”这一结果”，但事实上“词语运动特征”可能体现的仅仅是词语的运动特征，和词语的抽象性无关。换言之，运动特征可能只是具体的运动特征。例如，“本绪：把责任推出去”。被试评定其运动特征可能是基于“推出去”这个具体动作进行的。因此，抽象动词具身性的机制在于人们基于具体的运动体验来理解抽象的动作，符合概念隐喻理论。该研究探讨抽象动词具身性的神经机制这一问题很有价值，但不能仅仅停留于讨论相关脑区的层面，相关脑区的激活依然是一种描述性结果，未能真正解释机制。建议作者从概念隐喻、进化等角度进一步补充抽象动词具身性的机制（仅供参考，作者也可以选择其他更合适的理论视角）。

回应: 非常感谢审稿专家严谨而中肯的建议！审稿专家提出的“运动特征可能只是具体的运动特征。”...“抽象动词具身性的机制在于人们基于具体的运动体验来理解抽象的动作，符合概念隐喻理论。”这一问题，确实是我们在解释和讨论结果时忽略的内容。感谢审稿专家的建议，我们已将此部分的内容补充在了讨论部分，详见修改稿第 4 部分总讨论 4.2 小节的第三自然段。

意见 2: 文章篇幅较长，尤其是讨论部分，仅第一小点就用了 3500 字左右，请根据杂志对字数的要求进行精炼。

回应: 非常感谢审稿专家严谨而中肯的建议！笔者对总讨论部分的内容进行了反复修整，重新组织了部分前后重复的内容，并试图在不改变内容主旨的前提下更为简洁地描述内容，详见修改稿第 4 部分总讨论。另根据杂志对文章其他部分的字数要求，笔者对全文内容进行了校对和删减。

第三轮

审稿人 3 意见: 作者认真回答了二审问题并进行了修改, 无进一步问题。

编委意见: 同意发表。

主编意见: 同意外审和编委意见, 建议录用。